



Europäisches Patentamt  
European Patent Office  
Office européen des brevets

(11) Veröffentlichungsnummer:

**0 211 299**  
**A2**

(12)

## EUROPÄISCHE PATENTANMELDUNG

(21) Anmeldenummer: 86109945.5

(22) Anmeldetag: 19.07.86

(51) Int. Cl.<sup>4</sup>: **C 12 N 15/00**  
**C 07 K 13/00, C 12 P 21/02**  
**C 12 N 1/20**  
**/(C12N1/20, C12R1:19)**

(30) Priorität: 27.07.85 DE 3526995

(43) Veröffentlichungstag der Anmeldung:  
25.02.87 Patentblatt 87/9

(84) Benannte Vertragsstaaten:  
AT BE CH DE FR GB IT LI LU NL SE

(71) Anmelder: HOECHST AKTIENGESELLSCHAFT  
Postfach 80 03 20  
D-6230 Frankfurt am Main 80(DE)

(72) Erfinder: Habermann, Paul, Dr.  
Heimchenweg 80  
D-6230 Frankfurt am Main 80(DE)

(72) Erfinder: Stengelin, Siegfried, Dr. c/o Massachusetts  
General Hospital Dept. of Molecular Biology  
Fruit Street Boston, MA 02114(DE)

(72) Erfinder: Wengenmayer, Friedrich, Dr.  
Am Seyenbach 38  
D-6238 Hofheim am Taunus(DE)

(54) Fusionsproteine, Verfahren zu ihrer Herstellung und ihre Verwendung.

(57) Ein Abschnitt von etwa 70 Aminosäuren aus dem D-Protein des trp-Operons von E. coli eignet sich zur gentechnischen Konstruktion von Fusionsproteinen, die vor dem N-Terminus des gewünschten Proteins eine kurze Aminosäuresequenz aus genetisch kodierbaren Aminosäuren enthalten können, welche vorzugsweise Lys-Ala ist oder N-terminal Lys-Ala enthält.

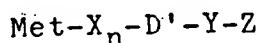
EP 0 211 299 A2

# Fusionsproteine, Verfahren zu ihrer Herstellung und ihre Verwendung

- Bei der gentechnischen Herstellung kleinerer eukaryotischer Proteine mit einem Molgewicht bis zu etwa 15 000 Dalton wird in Bakterien häufig nur eine geringe Ausbeute erhalten. Vermutlich werden die gebildeten Proteine durch wirtseigene Proteasen rasch abgebaut. Solche Proteine werden deshalb zweckmäßig als Fusionsproteine, insbesondere mit einem wirtseigenen Proteinanteil, hergestellt, der dann anschließend abgespalten wird.
- 10 Es wurde nun gefunden, daß ein aus nur etwa 70 Aminosäuren bestehender Abschnitt des D-Proteins aus dem trp-Operon von E. coli sich besonders gut zur Bildung von Fusionsproteinen eignet, und zwar im Bereich der Aminosäuresequenz 23 bis 93 (C. Yanofsky et al., Nucleic Acids Res. 9 (1981) 6647), im folgenden auch als "D'-Peptid" bezeichnet. Zwischen dem Carboxy-Terminus dieses Peptids und der Aminosäuresequenz des gewünschten eukaryotischen Proteins befindet sich eine Sequenz aus einer oder mehreren genetisch kodierbaren Aminosäuren, die eine chemische oder enzymatische
- 15 Abspaltung des gewünschten Proteins erlaubt. In bevorzugten Ausgestaltungen dieser Erfindung schließt sich an den Amino-terminus eine kurze Aminosäuresequenz an, bestehend aus Lys-Ala, gegebenenfalls gefolgt von einer Sequenz von 1 bis 10, insbesondere 1 bis 3 weiteren genetisch kodierbaren Aminosäuren, vorzugsweise von zwei Aminosäuren, insbesondere von Lys-Gly.
- 20
- 25

Die Erfindung betrifft somit ein Fusionsprotein der allgemeinen Formel

30



in der

n Null oder 1 ist,

BEST AVAILABLE COPY

- X eine Sequenz von 1 bis 12 genetisch kodierbaren Aminosäuren ist, vorzugsweise Lys-Ala,  
D' eine Sequenz von etwa 70 Aminosäuren im Bereich der Aminosäuresequenz 23-93 des D-Peptids im trp-Operon von E. coli ist,  
5 Y eine Sequenz aus einer oder mehreren genetisch kodierbaren Aminosäuren bedeutet, die eine Abspaltung der folgenden Aminosäuresequenz Z ermöglicht, und  
10 Z eine Sequenz aus genetisch kodierbaren Aminosäuren ist.

Weitere Aspekte der Erfindung und bevorzugte Ausgestaltungen werden im folgenden wiedergegeben bzw. in den  
15 Patentansprüchen definiert.

Es ist natürlich vorteilhaft, wenn der unerwünschte (wirts-eigene) Anteil am Fusionsprotein möglichst klein ist, da dann die Zelle nur wenig "Ballast" produziert und dadurch  
20 die Ausbeute an dem gewünschten Protein hoch ist. Außerdem entstehen bei der Abspaltung des unerwünschten Anteils weniger Nebenprodukte, was die Aufarbeitung erleichtert. Dem steht entgegen, daß die (vermutete) "Schutzfunktion" des unerwünschten Anteils nur ab einer bestimmten Größe  
25 zu erwarten ist. Überraschenderweise zeigte sich nun, daß das erfindungsgemäß gewählte Segment aus dem D-Protein diese Aufgabe erfüllt, obwohl es nur etwa 70 Aminosäuren aufweist.

30 In vielen Fällen, vor allem bei der bevorzugten Ausführungsform, in der X für Lys-Ala steht oder N-terminal diese Sequenz enthält, wird ein unlösliches Fusionsprotein gebildet. Dieses kann einfach von den löslichen Proteinen abgetrennt werden, was die Aufarbeitung sehr  
35 erleichtert und die Ausbeute erhöht. Die Bildung eines unlöslichen Fusionsproteins ist überraschend, da einerseits der bakterielle Anteil mit nur etwa 70 Aminosäuren recht gering ist und andererseits Bestandteil eines in der Wirtszelle in Lösung vorliegenden Proteins ist.

"Etwa 70 Aminosäuren im Bereich der Aminosäuresequenz 23 bis 93 des D-Peptids" besagt, daß in an sich bekannter Weise Variationen möglich sind, also einzelne Aminosäuren weggelassen, ersetzt oder ausgetauscht werden können, ohne  
5 daß hierdurch die erfindungsgemäßen Fusionsproteine in ihren Eigenschaften signifikant verändert würden. Solche Variationen sind ebenfalls Gegenstand der Erfindung.

Das gewünschte eukaryotische Protein ist bevorzugt ein  
10 biologisch aktives Protein wie ein Hirudin oder eine Vorstufe eines solchen Proteins wie humanes Proinsulin.

Das Fusionsprotein wird durch Expression in einem geeigneten System gewonnen und in der besonders bevorzugten  
15 Ausführungsform nach Aufschluß der Wirtszellen aus dem Niederschlag isoliert, in dem es sich aufgrund der Schwerlöslichkeit anreichert. Es ist somit eine leichte Abtrennung von den löslichen Bestandteilen der Zelle möglich.

20 Als Wirtszellen kommen alle in Betracht, für die Expressionssysteme bekannt sind, also Säugerzellen und Mikroorganismen, vorzugsweise Bakterien, insbesondere E. coli, da ja der bakterielle Anteil des Fusionsproteins  
25 ein wirtseigenes Protein von E. coli ist.

Die DNA-Sequenz, die für das erfindungsgemäße Fusionsprotein kodiert, wird in bekannter Weise in einen Vektor eingebaut, der in dem gewählten Expressionssystem eine gute  
30 Expression gewährleistet.

In bakteriellen Wirten wählt man zweckmäßig den Promotor und Operator aus der Gruppe Lac, Tac,  $P_L$  oder  $P_R$  des Phagen  $\lambda$ , hsp, omp oder einen synthetischen Promotor, wie  
35 sie beispielsweise in der deutschen Offenlegungsschrift 3 430 683 (Europäische Patentanmeldung O 173 149) beschrieben sind.

Besonders geeignet ist ein Vektor, der aus dem trp-Operon aus *E. coli* die folgenden Elemente enthält: den Promotor, den Operator und die Ribosomenbindungsstelle des L-Peptids. Im kodierenden Bereich schließen sich besonders zweckmäßig die ersten drei Aminosäuren dieses L-Peptids an, worauf dann über eine kurze Aminosäuresequenz die Aminosäuren 23 bis 93 des D-Proteins im trp-Operon folgen.

Die Zwischensequenz Y, die eine Abspaltung des gewünschten Polypeptids ermöglicht, richtet sich nach der Zusammensetzung dieses gewünschten Peptids: Wenn dieses beispielsweise kein Methionin enthält, kann Y Met bedeuten, worauf eine chemische Spaltung mit Bromcyan folgt. Steht im Bindeglied Y am Carboxyterminus Cystein oder steht Y für Cys, so kann eine enzymatische Cystein-spezifische Spaltung oder eine chemische Spaltung, beispielsweise nach spezifischer S-Cyanylierung, folgen. Steht im Brückenglied Y am Carboxyterminus Tryptophan oder steht Y für Trp, so kann eine chemische Spaltung mit N-Bromsuccinimid erfolgen. Steht Y für Asp-Pro, so kann in an sich bekannter Weise (D. Piskiewicz et al., Biochemical and Biophysical Research Communications 40 (1970) 1173-1178) eine proteolytische Spaltung erfolgen. Die Asp-Pro-Bindung kann, wie gefunden wurde, noch säurelabiler gestaltet werden, wenn Y

(Asp)<sub>m</sub>-Pro bzw. Glu-(Asp)<sub>m</sub>-Pro ist, wobei m 1, 2 oder 3 bedeutet. Hierbei werden Spaltprodukte erhalten, die N-terminal mit Pro beginnen bzw. C-terminal mit Asp enden.

Beispiele für enzymatische Spaltungen sind ebenfalls bekannt, wobei auch modifizierte Enzyme mit verbesserter Spezifität eingesetzt werden können (vgl. C.S. Craik et al., Science 228 (1985) 291-297). Ist das gewünschte eukaryotische Peptid humanes Proinsulin, so kann man als Sequenz Y eine Peptidsequenz wählen, bei der eine durch Trypsin abspaltbare Aminosäure (Arg, Lys) an die N-terminale Aminosäure (Phe) des Proinsulins gebunden ist, beispielsweise

weise Ala-Ser-Met-Thr-Arg, da dann die Arginin-spezifische Spaltung mit der Protease Trypsin erfolgen kann. Falls das gewünschte Protein nicht die Aminosäurefolge Ile-Glu-Gly-Arg enthält, ist auch eine Spaltung mit Faktor Xa möglich (EP-A O 161 937).

Bei der Ausgestaltung der Sequenz Y hat man auch die Möglichkeit, auf die synthetischen Gegebenheiten Rücksicht zu nehmen und geeignete Spaltstellen für Restriktionsenzyme einzubauen. Die der Aminosäuresequenz Y entsprechende DNA-Sequenz kann deshalb auch die Funktion eines Linkers oder Adapters übernehmen.

Das erfindungsgemäße Fusionsprotein wird vorteilhaft unter der Kontrolle des trp-Operons von E. coli exprimiert. Ein den Promotor und den Operator des trp-Operons enthaltender DNA-Abschnitt ist inzwischen handelsüblich. Die Expression von Proteinen unter der Kontrolle des trp-Operons ist vielfach beschrieben, beispielsweise in der EP-A2 O 036 776. Die Induktion des trp-Operons kann durch Abwesenheit von L-Tryptophan und/oder Anwesenheit von Indolyl-3-acrylsäure im Medium erreicht werden.

Unter der Kontrolle des trp-Operators erfolgt zunächst die Transkription des L-Peptids, das aus 14 Aminosäuren besteht. Dieses L-Peptid enthält in den Positionen 10 und 11 jeweils L-Tryptophan. Die Geschwindigkeit der Proteinsynthese des L-Peptid bestimmt, ob die nachfolgenden Strukturgene ebenfalls translatiert werden, oder ob die Proteinsynthese abgebrochen wird. Bei Mangel an L-Tryptophan erfolgt nun eine langsame Synthese des L-Peptids infolge einer geringen Konzentration an der tRNA für L-Tryptophan und es erfolgt eine Synthese der folgenden Proteine. Bei hohen Konzentrationen an L-Tryptophan wird dagegen der entsprechende Abschnitt der mRNA rasch abgelesen und es kommt zum Abbruch der Proteinbiosynthese, da die mRNA eine Terminator-ähnliche Struktur annimmt (C. Yanofsky et al., a.a.O.).

Die Häufigkeit der Translation einer mRNA wird stark von der Art der Nukleotide in der Umgebung des Startkodons beeinflusst. Bei der Expression eines Fusionsproteins mit Hilfe des trp-Operons erscheint es daher günstig, die Nukleotide für die ersten Aminosäuren des L-Peptids für den Beginn des Strukturgens des Fusionsproteins einzusetzen. Bei der bevorzugten Ausführung der Erfindung im trp-System wurden die Nukleotide der ersten drei Aminosäuren des L-Peptids (im folgenden L'-Peptid) als Kodons für die N-terminalen Aminosäuren des Fusionsproteins gewählt.

Die Erfindung betrifft deshalb auch Vektoren, vorzugsweise Plasmide zur Expression von Fusionsproteinen, wobei die DNA der Vektoren vom 5'-Ende aus die folgenden Merkmale (in geeigneter Anordnung und in Phase) besitzt: einen Promotor, einen Operator, eine Ribosomen-Bindungsstelle und das Strukturgen für das Fusionsprotein, wobei das letztere die Aminosäuresequenz I (Anhang) vor der Sequenz des gewünschten Proteins enthält. Vor dem Strukturgen bzw. als erstes Triplet des Strukturgens findet sich das Startkodon (ATG) und gegebenenfalls weitere Kodons für andere Aminosäuren, die zwischen dem Start-Kodon und der D'-Sequenz oder zwischen der D'-Sequenz und dem Gen für das gewünschte Protein angeordnet sind. Je nach der Aminosäurezusammensetzung des gewünschten Proteins wird die Auswahl der DNA-Sequenz vor dem Strukturgen getroffen, damit eine Abspaltung des gewünschten Proteins aus dem Fusionsprotein ermöglicht wird.

Bei der Expression des erfindungsgemäßen Fusionsproteins kann es sich als zweckmäßig erweisen, einzelne Triplets der ersten Aminosäuren nach dem ATG-Startkodon zu verändern, um eine eventuelle Basenpaarung auf der Ebene der mRNA zu verhindern. Solche Veränderungen, ebenso wie Veränderungen, Deletionen oder Additionen einzelner Aminosäuren im D'-Protein sind dem Fachmann geläufig und ebenfalls Gegenstand der Erfindung.

Da kleinere Plasmide mehrere Vorteile bieten, besteht eine bevorzugte Ausgestaltung der Erfindung darin, aus Plasmiden, die sich vom pBR 322 ableiten, einen DNA-Abschnitt mit dem Strukturgen für die Tetracyclinresistenz zu eliminieren. Vorteilhaft entfernt man das Segment von der HindIII-Schnittstelle bei Position 29 bis zur PvuII-Schnittstelle bei Position 2066. Besonders zweckmäßig ist es, bei den erfindungsgemäßen Plasmiden noch einen etwas größeren DNA-Abschnitt zu entfernen, indem man die PvuII-Schnittstelle am Anfang (in Ableserichtung) des trp-Operons nutzt (die in einem nichtessentiellen Teil liegt). Man kann so das entstehende große Fragment mit den beiden PvuII-Schnittstellen direkt ligieren. Das erhaltene, um etwa 2 kbp verkleinerte Plasmid bewirkt eine Expressionssteigerung, die möglicherweise auf eine erhöhte Kopienzahl in der Wirtszelle zurückzuführen ist.

In den folgenden Beispielen wird die Erfindung näher erläutert.

20

#### Beispiel 1

a) Chromosomale E. coli-DNA wird mit Hinf I geschnitten und das 492 bp-Fragment isoliert, das aus dem trp-Operon den Promotor, Operator, das Strukturgen des L-Peptids, den Attenuator und die Kodons für die ersten sechs Aminosäuren des trp-E-Strukturgens enthält. Dieses Fragment wird mit Hilfe von Klenow-Polymerase mit Desoxynukleotidtriphosphaten aufgefüllt, an beiden Enden mit einem Oligonukleotid verbunden, das eine Erkennungsstelle für HindIII enthält und mit HindIII nachgeschnitten. Das so erhaltene Hind III-Fragment wird in die HindIII-Schnittstelle von pBR 322 ligiert. Man erhält so das Plasmid ptrpE2-1 (J.C. Edmann et al., Nature 291 (1981) 503-506). Dieses wird wie beschrieben in das Plasmid ptrpL1 überführt.



Mit Hilfe der synthetischen Oligonukleotide (N1) und (N2)

5' CGA CAA TGA AAG CAA AGG 3' (N1)

5 5' CCT TTG CTT TCA TTG T 3' (N2)

die zu dem doppelsträngigen Oligonukleotid (N3)

10 5' CGA CAA TGA AAG CAA AGG 3'  
3' T GTT ACT TTC GTT TCC 5' (N3)

komplementieren, wird in die ClaI-Stelle des Plasmids ptrpL1 die DNA-Sequenz für die ersten drei Aminosäuren des L-Peptids eingebaut sowie eine Restriktionsstelle (StuI) für die Insertion weiterer DNA gebildet (Figur 1). Das Plasmid ptrpL1 wird mit dem Enzym ClaI nach den Angaben des Herstellers umgesetzt, die Mischung mit Phenol extrahiert und die DNA mit Ethanol ausgefällt. Das geöffnete Plasmid wird zur Entfernung der Phosphatgruppen an den 5'-Enden mit Alkalischer Phosphatase aus E. coli umgesetzt. Die synthetischen Nukleotide werden an den 5'-Enden phosphoryliert und mit T4 DNA-Ligase in das geöffnete, Phosphatase behandelte Plasmid eingesetzt. Nach beendeter Ligase-Reaktion erfolgt die Transformation in E. coli 294 und Selektion der Transformanten durch Amp-Resistenz und Anwesenheit einer StuI-Restriktionsstelle.

Etwa 80 % der erhaltenen Klone wiesen die erwartete Restriktionsstelle auf; die in Figur 1 wiedergegebene Nukleotidsequenz wurde durch Sequenzanalyse bestätigt. Man erhält das Plasmid pH120/14, das nach der Ribosomen-Bindungsstelle für das L-Peptid die Nukleotidtripletts für die ersten drei Aminosäuren des L-Peptids aufweist (L'-Peptid), gefolgt von einer StuI-Stelle, die ihrerseits die Insertion weiterer DNA erlaubt und so die Bildung von Fusionsproteinen mit den ersten drei Aminosäuren des L-Peptids ermöglicht.

b) Im folgenden wird am Beispiel des oben eingesetzten Oligonukleotids (N1) die chemische Synthese solcher Oligonukleotide erläutert:

5 Nach der Methode von M.J. Gait et al., Nucleic Acids  
Research 8 (1980) 1081-1096, wird das am 3'-Ende  
stehende Nukleosid, im vorliegenden Falle also Guanosin,  
kovalent an einen Glaskugel-Träger (CPG (= controlled  
10 pore glass) LCAA (= long chain alkyl amine) der Fa.  
Pierce) über die 3'-Hydroxyfunktion gebunden. Hierbei  
wird das Guanosin als N-2-Isobutyryl-3'-O-succinoyl-  
5'-dimethoxy-tritylether in Gegenwart von N,N'-Di-  
cyclohexylcarbodiimid und 4-Dimethylamino-pyridin mit  
dem modifizierten Träger umgesetzt, wobei die freie  
15 Carboxygruppe des Succinoylrestes den Aminorest des  
langkettigen Amins auf dem Träger acyliert.

In den folgenden Syntheseschritten wird die Basenkompo-  
nente als 5'-O-Dimethoxytrityl-nukleosid-3'-phosphorig-  
20 säuremonomethylester-dialkylamid oder -chlorid einge-  
setzt, wobei das Adenin als N6-Benzoyl-Verbindung, das  
Cytosin als N4-Benzoyl-Verbindung, das Guanin als  
N2-Isobutyryl-Verbindung und das keine Aminogruppe ent-  
haltende Thymin ohne Schutzgruppe vorliegen.

25 40 mg des Trägers, der 1 µmol Guanosin gebunden ent-  
hält, werden nacheinander mit den folgenden Agentien  
behandelt:

- 30 a) Methylenchlorid,  
b) 10 % Trichloressigsäure in Methylenchlorid,  
c) Methanol,  
d) Tetrahydrofuran,  
e) Acetonitril,  
35 f) 15 µmol des entsprechenden Nukleosidphosphits und  
70 µmol Tetrazol in 0,3 ml wasserfreiem Acetonitril  
(5 Minuten),

- g) 20 % Acetanhydrid in Tetrahydrofuran mit 40 % Lutidin und 10 % Dimethylaminopyridin (2 Minuten),  
h) Tetrahydrofuran,  
i) Tetrahydrofuran mit 20 % Wasser und 40 % Lutidin,  
5 j) 3 % Jod in Kollidin/Wasser/Tetrahydrofuran im Volumenverhältnis 5:4:1,  
k) Tetrahydrofuran und  
l) Methanol.

10 Unter "Phosphit" wird hierbei der Desoxyribose-3'-monophosphorigsäure-monomethylester verstanden, wobei die dritte Valenz durch Chlor oder eine tertiäre Aminogruppe, beispielsweise einen Diisopropylaminorest, abgesättigt ist. Die Ausbeuten der einzelnen Syntheseschritte können jeweils nach der Detritylierungsreaktion  
15 b) spektrophotometrisch durch Messung der Absorption des Dimethoxytritylkations bei einer Wellenlänge von 496 nm bestimmt werden.

20 Nach abgeschlossener Synthese des Oligonukleotids werden die Methylphosphatschutzgruppen des Oligomers mit Hilfe von p-Thiokresol und Triethylamin abgespalten.

25 Anschließend wird durch 3-stündige Behandlung mit Ammoniak das Oligonukleotid vom festen Träger abgetrennt. Eine 2- bis 3-tägige Behandlung der Oligomeren mit konzentriertem Ammoniak spaltet die Aminoschutzgruppen der Basen quantitativ ab. Das so erhaltene Rohprodukt wird durch Hochdruckflüssigkeitschromatographie (HPLC) oder durch Polyacrylamid-Gelelektrophorese  
30 gereinigt.

Ganz entsprechend werden auch die übrigen Oligonukleotide synthetisiert.

c) Das Plasmid ptrpE5-1 (R.A. Hallewell et al., Gene 9 (1980) 27-47) wird mit den Restriktionsenzymen HindIII und SalI nach den Angaben der Hersteller umgesetzt und das etwa 620 bp DNA-Fragment entfernt. Die synthetischen Oligonukleotide (N4) und (N5)

5' AGC TTC CAT GAC GCG T 3' (N4)

5' ACG CGT CAT GGA 3' (N5)

werden phosphoryliert, gemeinsam bei 37°C inkubiert und mit Hilfe von DNA-Ligase an die stumpfendige DNA-für Proinsulin (W. Wetekam et al., Gene 19 (1982) 179-183) addiert. Nach Umsetzen mit HindIII und SalI wird die nun verlängerte Proinsulin-DNA mit dem Enzym T4 DNA-Ligase kovalent in das geöffnete Plasmid eingebaut (Figur 2), wobei das Plasmid pH106/4 entsteht.

Das Plasmid pH106/4 wird zunächst nochmals mit SalI umgesetzt, die überlappenden Enden mit Klenow-Polymerase in stumpfe Enden ergänzt und anschließend mit dem Enzym MstI inkubiert. Es wird ein etwa 500 bp DNA-Fragment isoliert, das den gesamten kodierenden Teil des Proinsulins sowie einen etwa 210 pb großen Abschnitt des D-Proteins aus dem trp-Operon von E. coli enthält. Das DNA-Fragment ist stumpfendig und wird in die StuI-Stelle des Plasmids pH120/14 inseriert, wobei das Plasmid pH154/25 entsteht (Figur 3). Dieses ist zur Expression eines Fusionsproteins unter der Kontrolle des trp-Operons geeignet, bei dem nach dem L'- und das D'-Peptid die Aminosäuresequenz Ala-Ser-Met-Thr-Arg angeordnet ist, an die sich die Aminosäuresequenz des Proinsulins anschließt.

## 35 Beispiel 2

Das Plasmid pH154/25 (Figur 3) wird mit den Restriktionsenzymen BamHI und XmaIII umgesetzt. Die überstehenden

- Enden werden mittels Klenow-Polymerase aufgefüllt und mit T4 DNA-Ligase verknüpft. Man erhält das Plasmid pH254 (Figur 4), das sich zur Expression eines Fusionsproteins mit der Aminosäuresequenz L',D'-Proinsulin unter der Kontrolle des trp-Promotors eignet. Das Plasmid ist etwas kleiner als pH154/25, was von Vorteil sein kann.

### Beispiel 3

- 10 Durch Inkubation des Plasmids pH254 (Beispiel 2; Figur 4) mit den Restriktionsenzymen MluI und Sall wird ein DNA-Abschnitt mit 280 bp freigesetzt, der abgetrennt wird. Das Restplasmid wird mit Klenow-Polymerase in die stumpf-  
15 cyclisiert. Dadurch entsteht das Plasmid pH255 (Figur 4), das zur Insertion eines Strukturgens in eine der Restriktionsstellen MluI, Sall und EcoRI geeignet ist. Unter Induktionsbedingungen erfolgt die Bildung eines Fusions-  
20 proteins mit dem L',D'-Protein. Selbstverständlich können durch geeignete Linker weitere Restriktionsstellen in das Plasmid pH255 eingefügt werden.

### Beispiel 4

- 25 Das Plasmid pH154/25 (Figur 3) wird mit den Enzymen MluI und EcoRI inkubiert und das freigesetzte DNA-Fragment (etwa 300 bp) entfernt. Das Restplasmid wird mit Klenow-Polymerase aufgefüllt. Der Ringschluß erfolgt durch Einwirkung von DNA-Ligase. Das entstandene Plasmid pH256  
30 (Figur 5) kann zur Insertion von Strukturgenen in die EcoRI-Stelle verwendet werden.

### Beispiel 5

- 35 Durch Entfernen eines 600 bp-Fragments aus dem Plasmid pH256 (Beispiel 4; Figur 5) mit den Restriktionsenzymen

- BamHI und NruI erhält man das Plasmid pH257 (Figur 5).  
Hierzu wird pH256 zunächst mit BamHI inkubiert und stumpfe  
Enden mit Klenow-Polymerase erzeugt. Nach Inkubation mit  
NruI und Abtrennung des 600 bp-Fragments erfolgt die  
5 Bildung von pH257 nach Inkubation mit DNA-Ligase.

#### Beispiel 6

- Durch Insertion des lac-Repressors (P.J.Farabaugh, Nature  
10 274 (1978) 765-769) in das Plasmid pKK 177-3 (Amann et al.,  
Gene 25 (1983) 167) erhält man das Plasmid pJF118. Dieses  
wird mit EcoRI und SalI umgesetzt und das Restplasmid  
isoliert.
- 15 Aus dem Plasmid pH106/4 (Figur 2) wird durch Einwirkung  
von SalI und Inkubation mit MstI ein etwa 495 bp großes  
Fragment erhalten.

Die synthetisch gewonnenen Oligonukleotide (N6) und (N7)

20

5' ACG AAT TCA TGA AAG CAA AGG 3' (N6)

5' CCT TTG CTT TCA TGA ATT CGT 3' (N7)

- 25 werden phosphoryliert und mit DNA-Ligase an das stumpf-  
endige DNA-Fragment addiert. Durch Umsetzung mit EcoRI und  
SalI werden überlappende Enden freigesetzt, die eine  
Ligierung in das geöffnete Plasmid pJF118 erlauben.
- 30 Nach Transformation des so erhaltenen Hybridplasmids in  
E. coli 294 werden aufgrund der Größe der Restriktions-  
fragmente die richtigen Klone ausgewählt. Dieses Plasmid  
wird als pJ120 bezeichnet (Figur 6).
- 35 Die Expression des Fusionsproteins wird im Schüttelkolben  
wie folgt durchgeführt:

- Aus einer Übernachtskultur von E. coli 294-Transformanten, die das Plasmid pJ120 enthalten, in LB-Medium (J.H. Miller, Experiments in Molecular Genetics, Cold Spring Harbor Laboratory, 1972) mit 50 µg/ml Ampicillin wird eine frische
- 5 Kultur im Verhältnis von etwa 1:100 angesetzt und das Wachstum über OD-Messung verfolgt. Bei OD= 0,5 wird die Kultur mit soviel Isopropyl-β-D-galactopyranosid (IPTG) versetzt, daß dessen Konzentration 1 mM beträgt, und die Bakterien nach 150 bis 180 Minuten abzentrifugiert. Die
- 10 Bakterien werden 5 Minuten in einer Puffermischung (7M Harnstoff, 0,1 % SDS, 0,1 M Natriumphosphat, pH 7,0) gekocht und Proben auf eine SDS-Gelelektrophoreseplatte aufgetragen. Nach Elektrophorese wird aus Bakterien, die das Plasmid pJ120 enthalten, eine Proteinbande erhalten,
- 15 ten, die der Größe des erwarteten Fusionsproteins entspricht und mit Antikörpern gegen Insulin reagiert. Nach Isolierung des Fusionsproteins kann man durch Bromcyan-Spaltung das erwartete Proinsulin-Derivat freisetzen. Nach Aufschluß der Bakterien (French Press; <sup>(R)</sup>Dyno-Mühle) und
- 20 Zentrifugation befindet sich das L',D'-Proinsulin-Fusionsprotein im Niederschlag, so daß mit dem Überstand bereits erhebliche Mengen der übrigen Proteine abgetrennt werden können.
- 25 Die angegebenen Induktionsbedingungen gelten für Schüttelkulturen; bei größeren Fermentationen sind entsprechend veränderte OD-Werte und gegebenenfalls leicht variierte IPTG-Konzentrationen zweckmäßig.

30 Beispiel 7

- Aus E. coli 294-Transformanten, die das Plasmid pH154/25 (Figur 3) enthalten, wird eine Übernachtskultur in LB-Medium mit 50 µg/l Ampicillin hergestellt und am nächsten
- 35 Morgen im Verhältnis von etwa 1:100 in M9-Medium (J.H. Miller, a.a.O.) mit 2000 µg/ml Casamino Acids und 1 µg/ml

Auch im vorliegenden Falle gelten die angegebenen Induktionsbedingungen für Schüttelkulturen. Fermentationen in größerem Volumen erfordern veränderte Konzentrationen an Casamino Acids bzw. einen Zusatz an L-Tryptophan.

Das Plasmid pH154/25 (Figur 3) wird mit EcoRI geöffnet und  
20 die überstehenden DNA-Einzelstränge mit Klenow-Polymerase  
aufgefüllt. Die so erhaltene DNA wird mit dem Enzym MluI  
inkubiert und die für Insulin kodierende DNA aus dem  
Plasmid herausgeschnitten. Durch gelelektrophoretische  
Auftrennung wird dieses Fragment vom Restplasmid abge-  
25 trennt und das Restplasmid isoliert.

35

				Met	Thr			
5'	CCC	ACG	CGT	ATG	ACG	T	3'	
3'	GGG	TGC	GCA	TAC	TGC	ATA	5'	(N8)



- ligiert. Das Ligationsprodukt wird mit MluI inkubiert. Nach Inaktivierung des Enzyms bei 65°C wird das DNA-Gemisch mit Alkalischer Rinderphosphatase eine Stunde bei 37°C behandelt. Hierauf werden die Phosphatase und das Restriktionsenzym mittels Phenolextraktion aus dem Gemisch entfernt und die DNA durch Ethanolfällung gereinigt. Die so behandelte DNA wird mit T4-Ligase in das geöffnete Restplasmid pH154/25 eingesetzt, wobei das Plasmid pK150 erhalten wird, das durch Restriktionsanalyse und DNA-Sequenzierung nach Maxam und Gilbert charakterisiert wurde (Figur 7).

#### Beispiel 9

- E. coli 214-Bakterien, die das Plasmid pK150 (Figur 7) enthalten, werden in LB-Medium mit 30 bis 50 µg/ml Ampicillin über Nacht bei 37°C angezogen. Die Kultur wird mit M9-Medium, 2000 µg/l Casamino Acids und 1 µg/l Thiamin enthält, im Verhältnis 1:100 verdünnt und bei 37°C unter ständiger Durchmischung inkubiert. Bei einer  $OD_{600} = 0,5$  bzw. 1 wird bis zu einer Endkonzentration von 15 µg/l Indolyl-3-acrylsäure zugesetzt und 2 bis 3 Stunden bzw. 16 Stunden inkubiert. Anschließend werden die Bakterien abzentrifugiert und in 0,1 M Natriumphosphatpuffer (pH 6,5) unter Druck aufgeschlossen. Die schwerlöslichen Proteine werden abzentrifugiert und durch SDS-Polyacrylamid-Elektrophorese analysiert. Es zeigt sich, daß Zellen, deren trp-Operon induziert wurde, im Bereich unterhalb 20 000 Dalton, aber größer als 14 000 Dalton, ein neues Protein enthalten, das in nicht induzierten Zellen nicht gefunden wird. Nach Isolation des Fusionsproteins und Umsetzung mit Bromcyan wird Hirudin freigesetzt.

#### Beispiel 10

- Im folgenden werden Konstruktionen beschrieben, die es erlauben, vor das 5'-Ende der trp-D-Sequenz DNA-Sequenzen

einzuführen, die möglichst zahlreiche Erkennungsstellen für verschiedene Restriktionsenzyme enthalten, um die trp-D-Sequenz möglichst universell in die verschiedensten prokaryotischen Expressionssysteme einbauen zu können.

5

Die Plasmide pUC12 und pUC13 (Pharmacia P-L Biochemicals, 5401 St. Goar: The Molecular Biology Catalogue 1983, Appendix, p. 89) enthalten eine Polylinkersequenz, wobei im Plasmid pUC13 zwischen den Restriktionsschnittstellen für XmaI und SacI das MstI-HindIII-trp-Fragment aus dem Plasmid p106/4 (Figur 2), fusioniert mit dem HindIII-Hirudin-SacI-Fragment aus dem Plasmid pK150 (Figur 7) eingesetzt werden soll.

15 Hierzu wird die DNA des Plasmids pUC13 zunächst mit dem Restriktionsenzym XmaI behandelt. Die Enden des linearisierten Plasmids werden mittels Klenow-Polymerase-Reaktion aufgefüllt. Nach Ethanolfällung wird die DNA mit dem Enzym SacI behandelt und mit Ethanol erneut aus dem Reaktionsgemisch gefällt. Die DNA wird nun in einem wäßrigen Ligationsansatz mit dem MstI-HindIII-trp-D-Fragment, isoliert aus Plasmid p106/4, und dem HindIII-SacI-Hirudin-Fragment, isoliert aus dem Plasmid pK150, und T4 DNA-Ligase umgesetzt.

25

Das so erhaltene Plasmid pK160 enthält nun unmittelbar vor Beginn der trp-D-Sequenz eine multiple Restriktionsenzym-Erkennungssequenz, die Schnittstellen für die Enzyme XmaI, SmaI, BamHI, XbaI, HindII, SalI, AccI, PstI und HindIII umfaßt. Durch diese Konstruktion wird ferner nach dem 3'-Ende der Hirudin-Sequenz eine EcoRI-Schnittstelle erzeugt (Figur 8).

#### Beispiel 11

35

Das Plasmid p1131/5 wird (gemäß der nicht veröffentlichten deutschen Patentanmeldung P 35 14 113.1, Beispiel 1,

Figur 1) wie folgt hergestellt:

Das Plasmid ptrpL1 (J.C. Edman et al., Nature 291 (1981) 503-506) wird mit ClaI geöffnet und mit dem synthetisch hergestellten, selbst-komplementären Oligonucleotid (N9)

5        5'pCGACCATGGT 3';            (N9)  
ligiert.

Das so erhaltene Plasmid pH131/5 (Figur 9) wird an der  
derart eingeführten Schnittstelle für das Restriktions-  
10    enzym NcoI geöffnet und die entstandenen überstehenden  
Einzelstrangenden mittels Klenow-Polymerase-Reaktion aufge-  
füllt. Die linearisierte glattendige DNA wird nun mit dem  
Enzym EcoRI nachgeschnitten und die größere der beiden  
entstanden DNA-Sequenzen mittels Ethanolfällung von der  
15    kleineren Sequenz abgetrennt. Die so gewonnene Restplasmid-  
DNA des Plasmids pH131/5 wird nun mit einem für  
trp-D'-Hirudin kodierenden Fragment aus pK160 durch T4-  
Ligase-Reaktion ligiert. Dieses Fragment wurde durch Öffnen  
des Plasmids mit HincII und EcoRI aus dem Plasmid pK160  
20    herausgespalten. Das Fragment wird gelelektrophoretisch  
vom Restplasmid abgetrennt und anschließend aus dem Gel-  
material eluiert. Das Ligationsprodukt wird nach E. coli  
K12 transformiert. Die Plasmid-DNA enthaltenden Klone wer-  
den isoliert und durch Restriktionsanalyse und DNA-Sequenz-  
25    Analyse charakterisiert. Das so erhaltene Plasmid pK170  
enthält an den trp-Operator fusioniert eine DNA-Sequenz,  
die für Met-Asp-Ser-Arg-Gly-Ser-Pro-Gly-trp-D'-(Hirudin)  
kodiert (Figur 9).

### 30    Beispiel 12

Das Plasmid pJF118 (Beispiel 6) wird mit EcoRI geöffnet  
und die überstehenden DNA-Enden mittels Klenow-Polymerase-  
Reaktion in glatte Enden überführt. Die so behandelte  
DNA wird anschließend mit dem Enzym SalI geschnitten und  
35    gelelektrophoretisch von dem kurzen EcoRI SalI-Fragment  
abgetrennt.

Das Plasmid pK 170 (Beispiel 11) wird mit NcoI gespalten und die überstehenden Enden mittels Klenow-Polymerase in glatte Enden überführt. Die Plasmid-DNA wird aus dem Reaktionsgemisch durch Ethanolfällung abgetrennt und mit  
5 den Enzymen HindIII und BamHI behandelt. Von den entstehenden Fragmenten werden zwei isoliert, nämlich das NcoI (mit aufgefülltem Ende)-trpD'-HindIII-Fragment und das HindIII-Hirudin-BamHI-Fragment (deutsche Offenlegungsschrift 3 429 430). Beide Fragmente werden nach gelelektrophoretischer Auftrennung isoliert.  
10

Ferner wird das BamHI-Sall-HirudinII-Fragment gemäß Figur 2 der deutschen Offenlegungsschrift 3 429 430 isoliert. In einer Ligationsreaktion werden nun die vier Fragmente, nämlich das pJF 118-Restplasmid, das NcoI-trpD'-HindIII-Fragment, das HindIII-Hirudin-BamHI-Fragment und das  
15 HirudinII-Fragment, miteinander umgesetzt und das erhaltene Plasmid pK 180 (Figur 10) nach E. coli K12-W 3110 transformiert. Richtige Plasmide sind dadurch gegeben, daß ein EcoRI-trpD'-Hirudin-Sall-Fragment in der Plasmid-DNA nachgewiesen werden kann. Die trpD'-Hirudin-Sequenz ist nun an den tac-Promotor angeschlossen. Das Fusionsprotein wird  
20 gemäß Beispiel 6 exprimiert.

### 25 Beispiel 13

Bei den Plasmiden, die sich von pBR 322 ableiten, wie pH120/14 (Beispiel 1, Figur 1), pH154/25 (Beispiel 1, Figur 3), pH256 (Beispiel 4, Figur 5), pK150 (Beispiel 8, Figur 7) und pK170 (Beispiel 11, Figur 9) befindet sich -  
30 auf den Figuren im Uhrzeigersinn - zwischen dem Startcodon des Fusionsproteins und der nächsten HindIII-Stelle (entsprechend HindIII bei Position 29 in pBR322) eine zusätzliche PvuII-Stelle im Bereich des Fragments, das den trp-Promotor und -Operator enthält, allerdings außerhalb des Promotor-Bereichs.  
35

Es wurde nun festgestellt, daß durch Entfernen des DNA-Abschnittes, der durch die beschriebene PvuII-Stelle und die PvuII-Stelle, die der Position 2066 in pBR322 entspricht, begrenzt wird, die Ausbeute eines klonierten Proteins (oder Fusionsproteins) deutlich erhöht wird.

Im folgenden wird beispielhaft die Verkürzung des Plasmids pH154/25 zu pH154/25\* beschrieben, die für die übrigen obengenannten Plasmide entsprechend erfolgen kann (wobei die verkürzten Plasmide ebenfalls durch einen Stern gekennzeichnet sind):

pH154/25 wird (nach den Angaben des Herstellers) mit PvuII umgesetzt, wobei drei Fragmente entstehen:

- Fragment 1: Von der PvuII-Restriktionsstelle des Proinsulins bis zu der der Position 2066 in pBR322 entsprechenden PvuII-Restriktionsstelle,
- Fragment 2: von der PvuII-Restriktionsstelle nahe dem trp-Promotor bis zur PvuII-Stelle des Proinsulins und
- Fragment 3: von der PvuII-Stelle nahe dem trp-Promotor-Fragmentes bis zu der der Position 2066 in pBR322 entsprechenden PvuII-Stelle.

Die Fragmente können durch Elektrophorese auf Agarose getrennt und anschließend isoliert werden (Maniatis, a.a.O.).

- Die Fragmente 1 und 2 werden unter "blunt end"-Bedingungen mit dem Enzym T4 DNA-Ligase verbunden. Nach Transformation in E. coli 294 wird geprüft, in welchen Kolonien ein Plasmid mit der vollständigen Proinsulinsequenz vorhanden ist und somit die Fragmente in der gewünschten Anordnung vorliegen. Das Plasmid pH154/25\* ist in der Figur 11 dargestellt.

Bei der Expression, die wie in den vorstehenden Beispielen beschrieben erfolgt, wird eine deutliche Steigerung des Fusionsproteinanteils beobachtet.

15

#### Beispiel 14

Das Plasmid pH 154/25\* (Beispiel 13, Figur 11) wird mit HindIII und SalI verdaut und das kleine Fragment (mit der Proinsulinsequenz) gelelektrophoretisch abgetrennt. Das große Fragment wird isoliert und mit der synthetischen DNA (N10)

	(Ala)	Trp	Glu	Asp	Pro	Met	Ile	Glu	(Gly)	(Arg)	
A	GCT	TGG	GAG	GAT	CCT	ATG	ATC	GAG	GG		(N10)
		ACC	CTC	CTA	GGA	TAC	TAG	CTC	CCA	GCT	

ligiert. Es entsteht das Plasmid pInt13 (Figur 12).

Die DNA (N10) kodiert für eine Aminosäuresequenz, die mehrere Spaltstellen für eine chemische Spaltung enthält:

- a) Met für Bromcyan,
- b) Trp für N-Bromsuccinimid (NBS oder BSI)
- c) Asp-Pro für proteolytische Spaltung, wobei das vorangestellte Glu die Asp-Pro-Bindung zusätzlich gegenüber der Einwirkung von Säuren schwächt.

Die Einführung dieses HindIII-SalI-Linkers (N10) in den Leserahmen eines kodierten Polypeptids erlaubt also die angeführten Möglichkeiten für eine chemische Abspaltung, je nach der Aminosäurefolge des erwünschten Proteins bzw. nach seiner Empfindlichkeit gegenüber den Spaltungsagenzien.

Die Figuren sind nicht maßstabgerecht.

Aminosäuresequenz I

23

Ser Asn Gly His Asn Val Val Ile

10

Tyr Arg Asn His Ile Pro Ala Gln

20

Thr Leu Ile Glu Arg Leu Ala Thr

30

Met Ser Asn Pro Val Leu Met Leu

40

Ser Pro Gly Pro Gly Val Pro Ser

Glu Ala Gly Cys Met Pro Glu Leu

50

Leu Thr Arg Leu Arg Gly Lys Leu

60

Pro Ile Ile Gly Ile Cys Leu Gly

70

(93)

His Gln Ala Ile Val Glu Ala

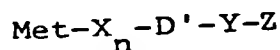
DNA-Sequenz I

Ser	Asn	Gly	His	Asn	Val	Val	Ile
AGC	AAT	GGG	CAT	AAC	GTG	GTG	ATT
			.			.	
	10						
Tyr	Arg	Asn	His	Ile	Pro	Ala	Gln
TAC	CGC	AAC	CAT	ATA	CCG	GCG	CAA
	.				.		
			20				
Thr	Leu	Ile	Glu	Arg	Leu	Ala	Thr
ACC	TTA	ATT	GAA	CGC	TTG	GCG	ACC
50			.				.
					30		
Met	Ser	Asn	Pro	Val	Leu	Met	Leu
ATG	AGT	AAT	CCG	GTG	CTG	ATG	CTT
		.			.		
						40	
Ser	Pro	Gly	Pro	Gly	Val	Pro	Ser
TCT	CCT	GGC	CCC	GGT	GTG	CCG	AGC
	100			.			.
Glu	Ala	Gly	Cys	Met	Pro	Glu	Leu
GAA	GCC	GGT	TGT	ATG	CCG	GAA	CTC
			.			.	
	50						
Leu	Thr	Arg	Leu	Arg	Gly	Lys	Leu
CTC	ACC	CGC	TTG	CGT	GGC	AAG	CTG
	150						
			60				
Pro	Ile	Ile	Gly	Ile	Cys	Leu	Gly
CCC	ATT	ATT	GGC	ATT	TGC	CTC	GGA
			.				.
					70		
His	Gln	Ala	Ile	Val	Glu	Ala	
Cat	CAG	GCG	ATT	GTC	GAA	GCT	
		200			.		
		.					



PATENTANSPRÜCHE:

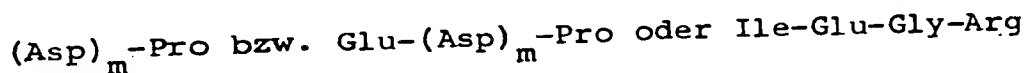
1. Fusionsprotein der allgemeinen Formel



- 5 in der  
n Null oder 1 ist,  
X eine Sequenz von 1 bis 12 genetisch kodierbaren  
Aminosäuren ist;  
D' eine Sequenz von etwa 70 Aminosäuren im Bereich  
10 der Aminosäuresequenz 23-93 des D-Peptids im trp  
Operon von E. coli ist,  
Y eine Sequenz aus einer oder mehreren genetisch  
kodierbaren Aminosäuren bedeutet, die eine Ab-  
spaltung der folgenden Aminosäuresequenz Z ermög-  
15 licht, und  
Z eine Sequenz aus genetisch kodierbaren Aminosäuren  
ist.

- 20 2. Fusionsprotein nach Anspruch 1, dadurch gekennzeichnet,  
daß n eins ist und X bis zu 5 Aminosäuren umfaßt, wobei  
vorzugsweise in X N-terminal Lys-Ala steht.

- 25 3. Fusionsprotein nach Anspruch 1 oder 2, dadurch gekenn-  
zeichnet, daß Y C-terminal Met, Cys, Trp, Arg oder Lys  
oder eine der Gruppen



- 30 in denen m 1, 2 oder 3 bedeutet,  
enthält oder aus diesen Aminosäuren oder einer dieser  
Gruppen besteht.

4. Fusionsprotein nach einem oder mehreren der vorher-  
gehenden Ansprüche, dadurch gekennzeichnet, daß Z die  
Aminosäuresequenz von humanem Proinsulin oder einem  
Hirudin bedeutet.

5. Verfahren zur Herstellung der Fusionsproteine nach einem oder mehreren der Ansprüche 1 bis 4, dadurch gekennzeichnet, daß man eine für diese Fusionsproteine kodierende Genstruktur in einer Wirtszelle, vorzugsweise in einem Bakterium, insbesondere E.coli, exprimiert und das Fusionsprotein abtrennt.
6. Verfahren nach Anspruch 5, dadurch gekennzeichnet, daß in der Genstruktur für D' die DNA-Sequenz I (Anhang) kodiert, und/oder in der Genstruktur für X die DNA-Sequenz (kodierender Strang)

15

5' AAA GCA AAG GGC 3'

kodiert, und/oder die Genstruktur so gewählt wird, daß das Fusionsprotein unlöslich ist, und/oder die Genstruktur in Phase in einem Vektor enthalten ist, der aus dem trp-Operon von E. coli den Promotor, den Operator und die Ribosomenbindungsstelle des L-Peptids enthält.

- 30 7. Verfahren nach Anspruch 6, dadurch gekennzeichnet, daß der Vektor ein Derivat von pBR322 ist, wobei aus der pBR322-DNA der Abschnitt von der HindIII-Stelle bei Position 29 bis zur PvuII-Stelle bei Position 2066 entfernt ist.
8. Genstruktur, kodierend für Fusionsproteine nach Anspruch 1 bis 4.
9. Vektor, vorzugsweise Derivat des Plasmids pBR322, wobei aus der pBR322-DNA der Abschnitt von der HindIII-Stelle bei Position 29 bis zur PvuII-Stelle bei Position 2066 entfernt ist, enthaltend eine Genstruktur nach Anspruch 8.

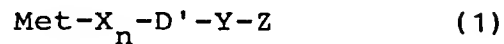
10. Expressionssystem, vorzugsweise E.coli-System, enthaltend einen Vektor nach Anspruch 9. 0211299

11. Verfahren zur Herstellung eines eukaryotischen Proteins, dadurch gekennzeichnet, daß man aus einem Fusionsprotein nach Anspruch 1 bis 4 die Aminosäuresequenz Z chemisch oder enzymatisch abspaltet.

12. Plasmide pH 154/25, pH 254, pH 255, pH 256, pH 257, pH 120/14, pK 150, pK 160, pK 170, pK 180, pH 154/25\*, pH 256\*, pH 120/14\*, pK 150\*, pK 170\* und pInt 13.

Patentansprüche Österreich:

1. Verfahren zur Herstellung eines Fusionsproteins der allgemeinen Formel (1)



in der

n Null oder 1 ist,

X eine Sequenz von 1 bis 12 genetisch kodierbaren Aminosäuren ist,

D' eine Sequenz von etwa 70 Aminosäuren im Bereich der Aminosäuresequenz 23-93 des D-Peptids im trp Operon von E. coli ist,

Y eine Sequenz aus einer oder mehreren genetisch kodierbaren Aminosäuren bedeutet, die eine Abspaltung der folgenden Aminosäuresequenz Z ermöglicht, und

Z eine Sequenz aus genetisch kodierbaren Aminosäuren ist, dadurch gekennzeichnet, daß man eine für diese Fusionsproteine kodierende Genstruktur in einer Wirtszelle exprimiert und das Fusionsprotein abtrennt.

2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß man eine Genstruktur für ein Fusionsprotein der Formel 1 einsetzt, in der n eins ist und X bis zu 5 Aminosäuren umfaßt.
3. Verfahren nach Anspruch 1 oder 2, dadurch gekennzeichnet, daß man eine Genstruktur für ein Fusionsprotein der Formel 1 einsetzt, in der n eins ist und in X N-terminal Lys-Ala steht.
4. Verfahren nach einem oder mehreren der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß man eine Genstruktur für ein Fusionsprotein der Formel 1 einsetzt,

in der Y C-terminal Met, Cys, Trp, Arg oder Lys oder eine der Gruppen (Asp)<sub>m</sub>-Pro bzw. Glu-(Asp)<sub>m</sub>-Pro oder Ile-Glu-Gly-Arg, in denen m 1, 2 oder 3 bedeutet, enthält oder aus diesen Aminosäuren besteht.

5. Verfahren nach einem oder mehreren der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß man eine Genstruktur für ein Fusionsprotein der Formel 1 einsetzt, in der Z die Aminosäuresequenz von humanem Proinsulin oder einem Hirudin bedeutet.
6. Verfahren nach einem oder mehreren der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß man eine Genstruktur für ein Fusionsprotein der Formel 1 einsetzt, in der für D' die DNA-Sequenz I (Anhang) kodiert.
7. Verfahren nach Anspruch 6, dadurch gekennzeichnet, daß in der Genstruktur für X die DNA-Sequenz (kodierender Strang)

5' AAA GCA AAG GGC 3'

kodiert.

8. Verfahren nach einem oder mehreren der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß die Genstruktur so gewählt wird, daß das Fusionsprotein unlöslich ist.
9. Verfahren nach einem oder mehreren der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß die Genstruktur in Phase in einem Vektor enthalten ist, der aus dem trp-Operon von E.coli den Promotor, den Operator und die Ribosomenbindungsstelle des L-Peptids enthält.
10. Verfahren nach Anspruch 9, dadurch gekennzeichnet, daß der Vektor ein Derivat von pBR322 ist, wobei aus der pBR322-DNA der Abschnitt von der HindIII-Stelle bei Position 29 bis zur PvuII-Stelle bei Position 2066 entfernt ist.

11. Verfahren nach einem oder mehreren der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß die Wirtszelle E.coli ist.
12. Verfahren zur Herstellung eines eukaryotischen Proteins, dadurch gekennzeichnet, daß man aus einem nach Anspruch 1 bis 5 erhaltenen Fusionsprotein die Aminosäuresequenz Z chemisch oder enzymatisch abspaltet.

FIG. 1

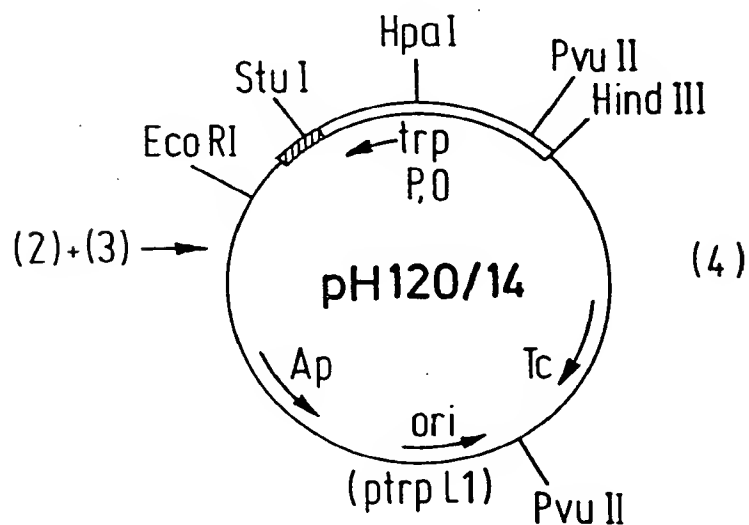
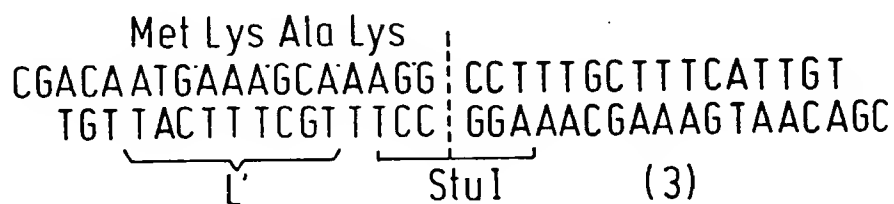
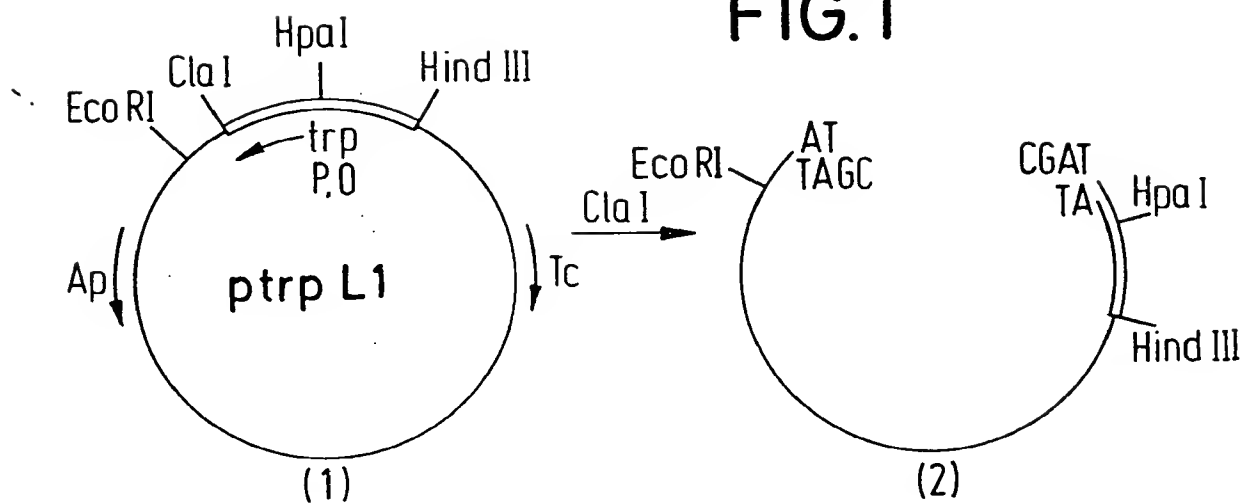


FIG. 2

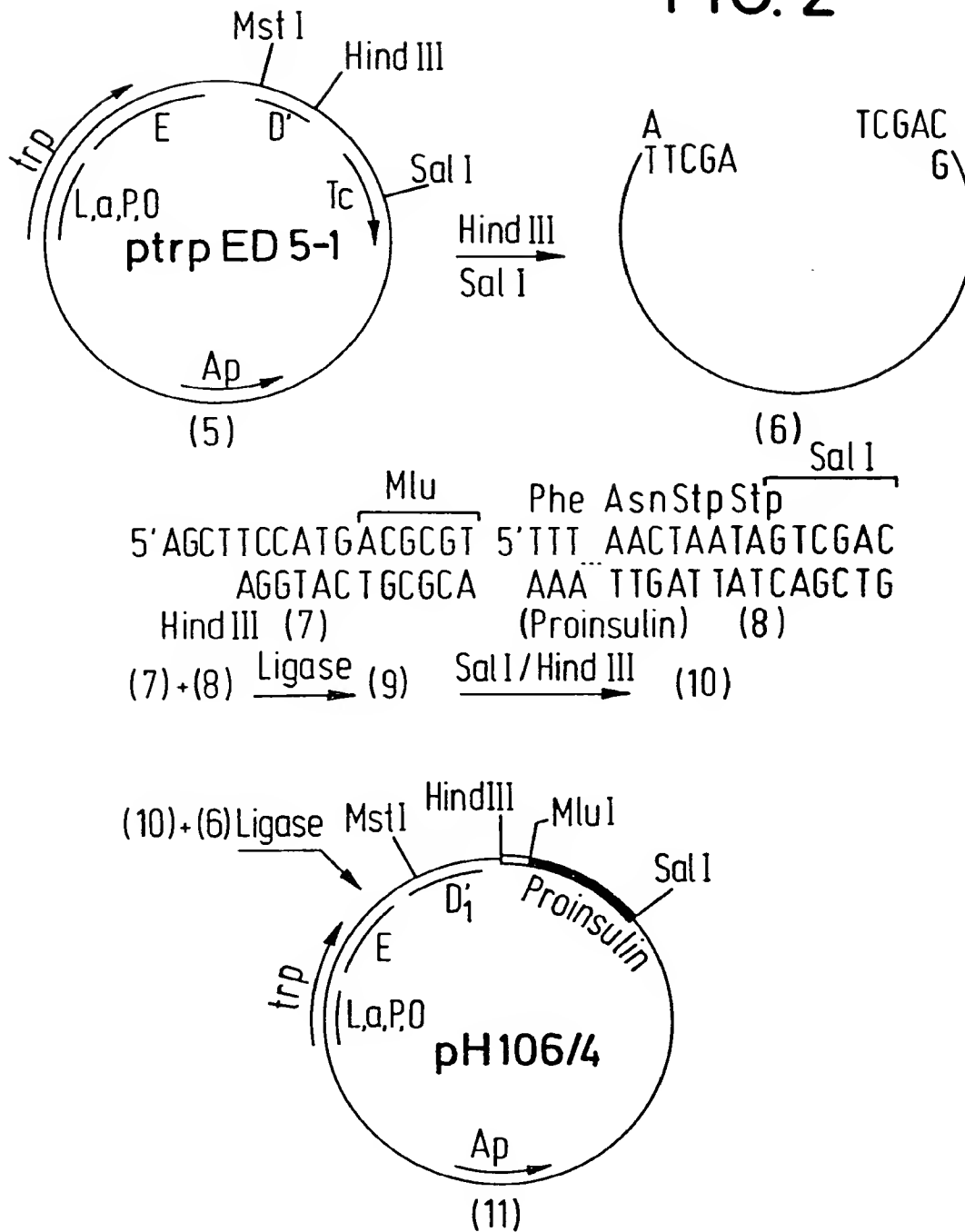
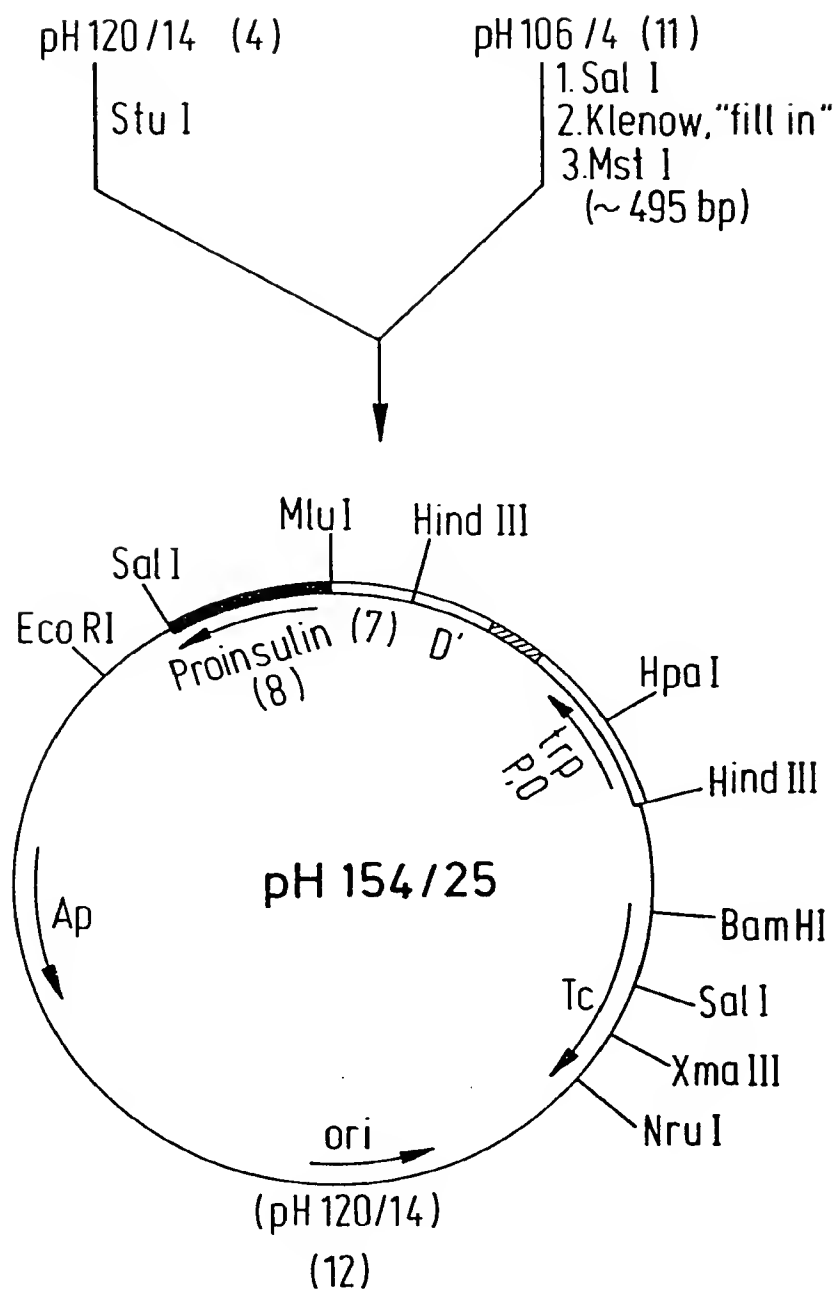
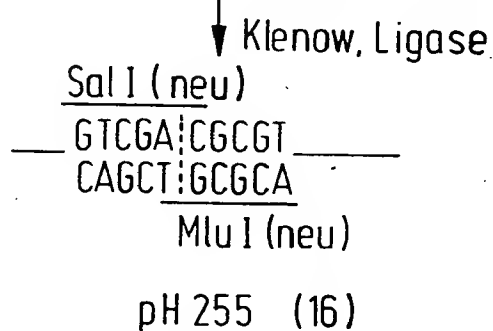
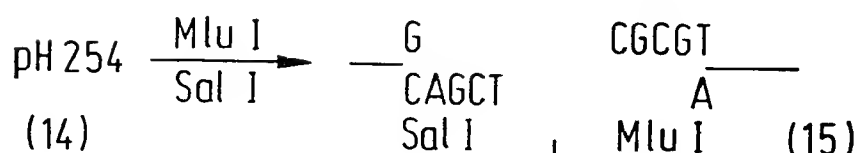
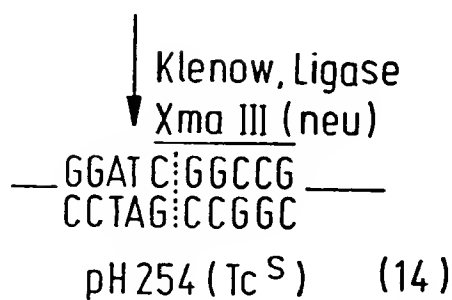
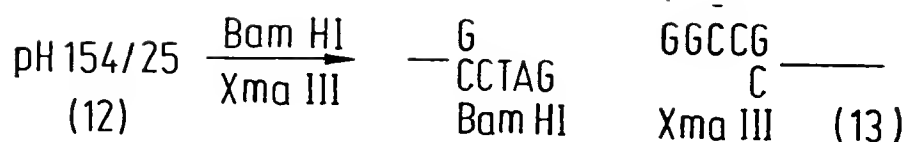




FIG. 3



## FIG. 4



## FIG. 5

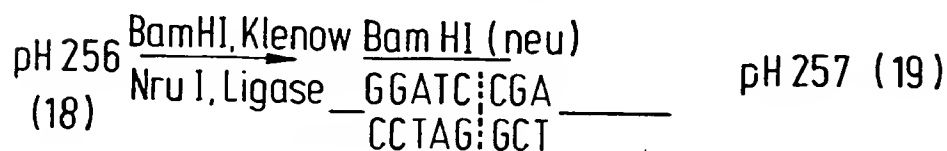
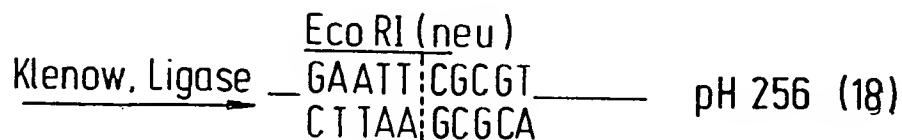
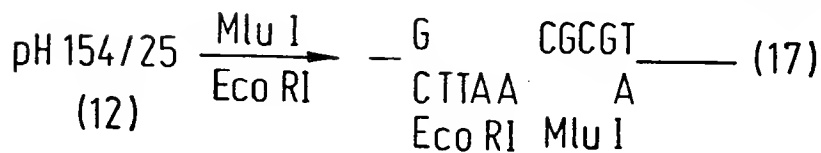
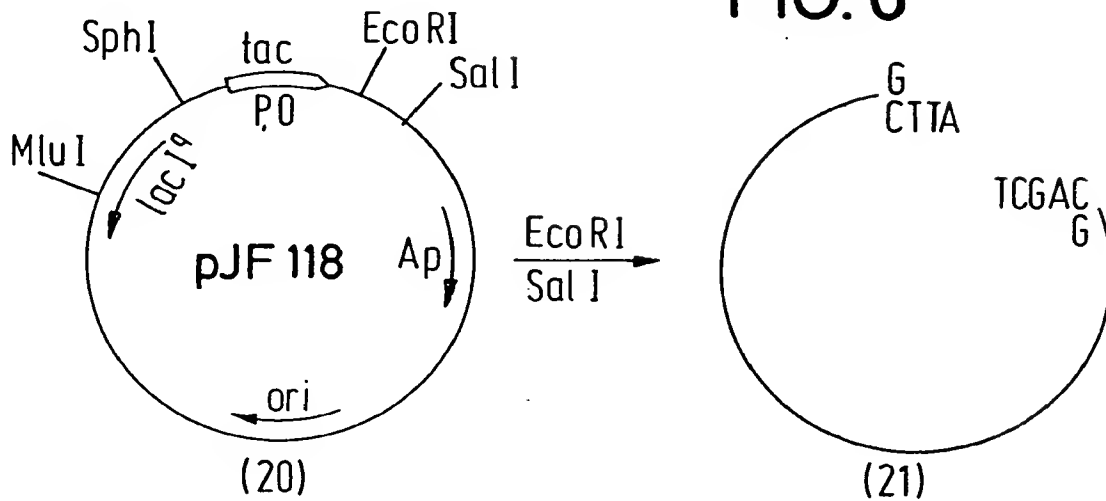


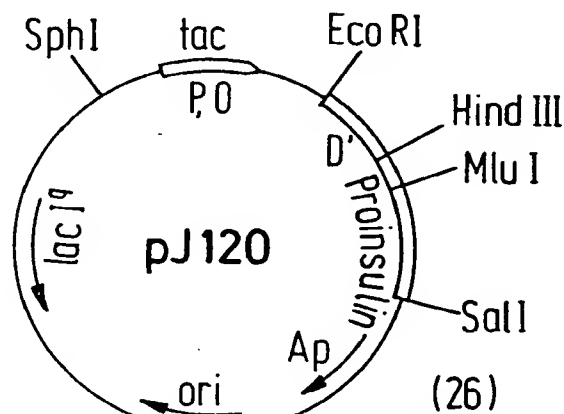
FIG. 6



pH 106/4  $\xrightarrow[\text{Sal I}]{\text{Mst I}}$  ~ 495 bp - Fragment (22)  
(11)

Met Lys Ala Lys  
5' AC GAA TTC ATG AAA GCA AAG G 3'  
3' TG CTT AAG TAC TTT CGT TTC C 5'  
(23)

(22) + (23)  $\xrightarrow{\text{Ligase}}$  (24)  $\xrightarrow[\text{Sal I}]{\text{Eco RI}}$  (25)  
(25) + (21)  $\xrightarrow{\text{Ligase}}$



Ausschnitt aus pJ120  
SD EcoRI Start Lys Ala Lys  
5'...CACAGGAAACAGAATTCATGAAA'GCA'AAG... 3'  
3'...GTGTCCTTTGTCTTAAGTAC TTT CGT TTC... 5'

FIG. 7

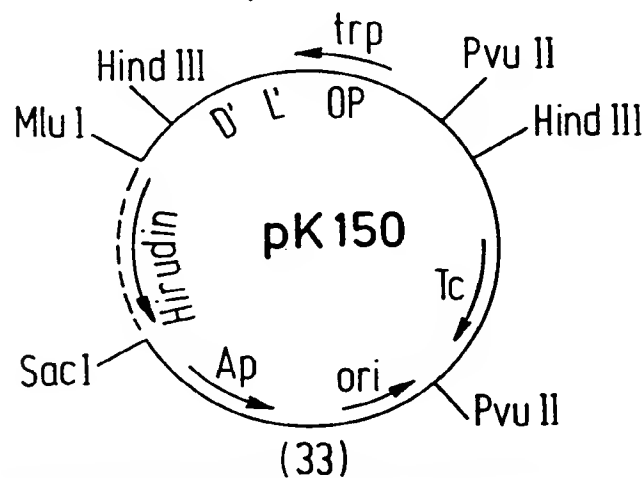
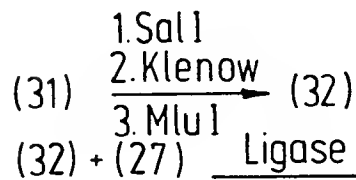
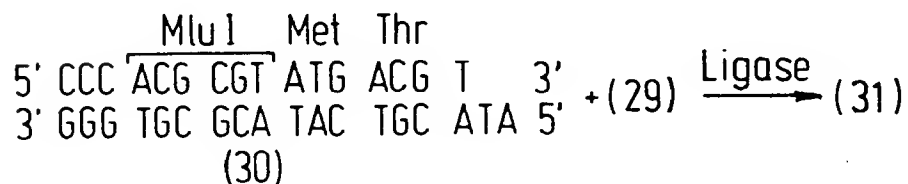
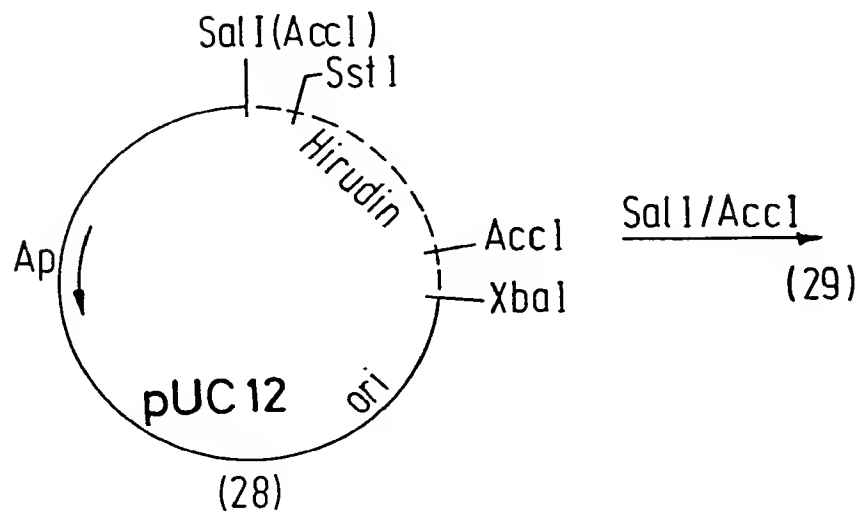
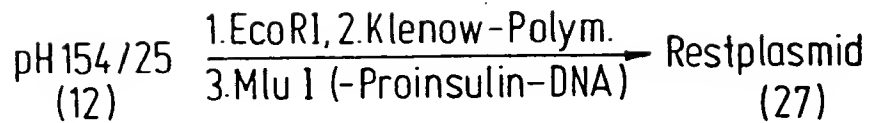


FIG. 8

pUC 13 (34)  $\xrightarrow[3. \text{ Sac I}]{1. \text{ Xma I}, 2. \text{ Klenow-P}}$  großes Fragment (35)

pK 150 (33)  $\xrightarrow[\text{Sac I}]{\text{Hind III}}$  Hind III - Hirudin - Sac I (36)

pH 106/4 (11)  $\xrightarrow[\text{Hind III}]{\text{Mst I}}$  Mst I - trp D' - Hind III (37)

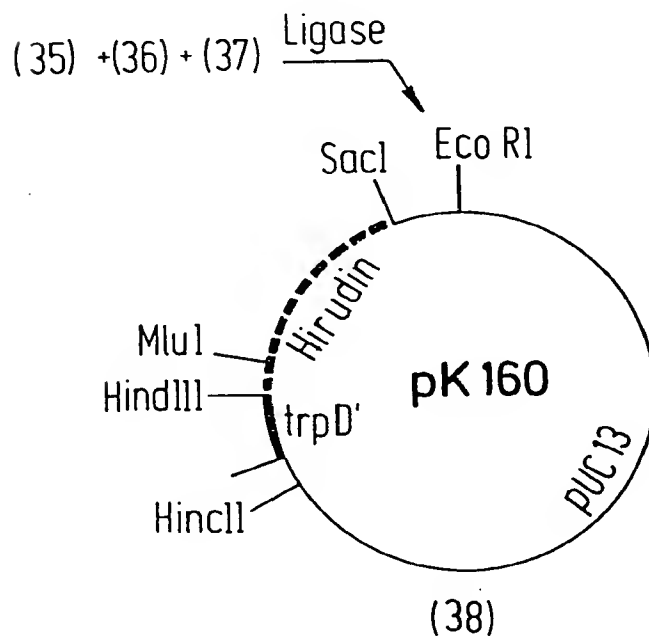


FIG. 9

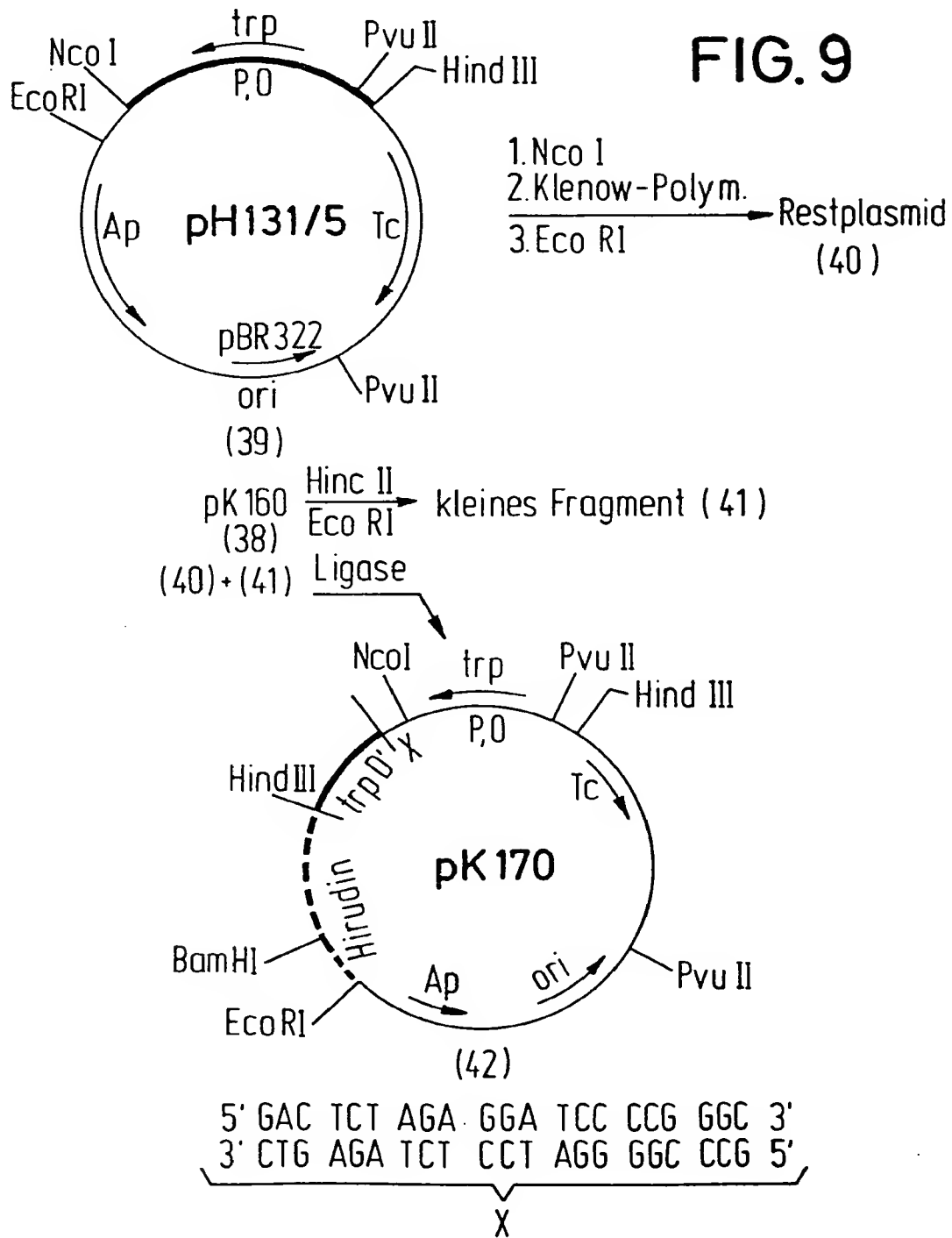
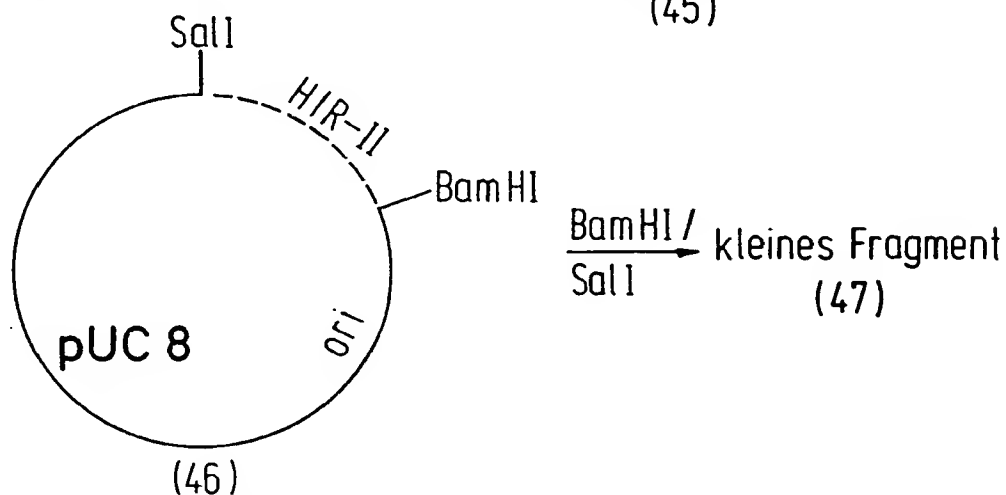


FIG.10

pJF 118 (20)  $\xrightarrow[3. \text{Sal I}]{1. \text{Eco RI} \atop 2. \text{Klenow-P}}$  Restplasmid (43)

pK170 (42)  $\xrightarrow[3. \text{Hind III/Bam HI}]{1. \text{Nco I} \atop 2. \text{Klenow-P}}$  NcoI<sup>-</sup> - trpD' - Hind III (44)  
+  
Hind III - Hirudin - Bam HI (45)



(43) + (44) + (45) + (47)  $\xrightarrow{\text{Ligase}}$

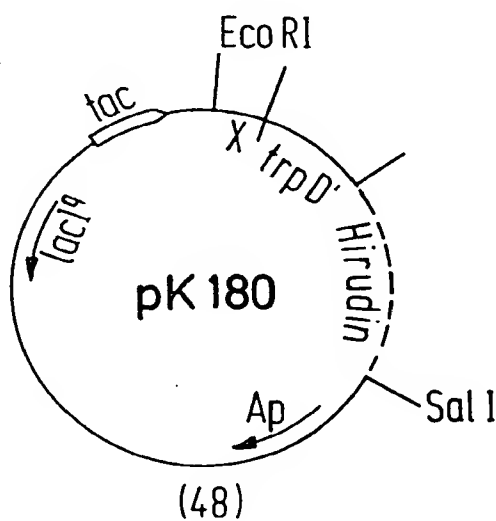


Diagram illustrating the construction of pH 154/25\* from pH 154/25.

**pH 154/25 (Top):** A circular plasmid map showing three fragments:

- Fragment 1 (12):** Contains the *Proinsulin* gene (8) and the *Ap* (Ampicillin resistance) gene.
- Fragment 2 (7):** Contains the *D'* region.
- Fragment 3 (14):** Contains the *ori* (pH 120/14) (origin of replication), *Tc* (Tetracycline resistance), and various restriction sites (*Sal*I, *Xma*III, *Nru*I, *Bam*HI, *Hind*III, *Hpa*I, *Pvu*II).

The plasmid also contains a *trp* P,0 region and a *Pvu*II site.

**Construction:** Fragments 1 (49), 2 (50), and 3 (49) + (50) are ligated using *Ligase*.

**pH 154/25\* (Bottom):** The resulting plasmid map, labeled (51), shows the assembled structure with the *ori* (pH 120/14) and *Proinsulin* (8) gene.



FIG.12

(51)  $\xrightarrow[\text{Sal I}]{\text{Hind III}}$  (10) + Restplasmid (52)

	(Ala)	Trp	Glu	Asp	Pro	Met	Ile	Glu	(Gly)	(Arg)
A	GCT	TGG	GAG	GAT	CCT	ATG	ATC	GAG	GG	
		ACC	CTC	CTA	GGA	TAC	TAG	CTC	CCA	GCT
	(Hind III)			Bam HI						(Sal I)

(53)

